HOUSEHOLD SOCIO-ENVIRONMENT AND INFANT FECAL MICROBIOME

By

MARYAM EDRISI

A thesis submitted in partial fulfillment of
the requirements for the degree of

MASTER OF ARTS IN ANTHROPOLOGY

WASHINGTON STATE UNIVERSITY
Department of Anthropology

MAY 2024

To the Faculty of Washington State University:

The members of the Committee appointed to examine the thesis of MARYAM EDRISI find it satisfactory and recommend that it be accepted.

Aaron Blackwell, Ph.D., Chair

Courtney L. Meehan, Ph.D.

Edward H. Hagen, Ph.D.

Elizabeth A. Holdsworth, Ph.D.

ACKNOWLEDGMENT

This research was supported by the Health Equity Research Center (HERC) at Washington State University (awarded to Courtney L. Meehan, Michelle K. McGuire, and Maria Gartstein), the USDA National Institute of Food and Agriculture (Hatch project #1020084 awarded to Michelle K. McGuire), and National Institutes of Health #P30GM103324. The funders had no role in study design, data collection and analysis, or preparation of the manuscript.

HOUSEHOLD SOCIO-ENVIRONMENT AND INFANT FECAL MICROBIOME

Abstract

by Maryam Edrisi, M.A.
Washington State University
May 2024

Chair: Aaron Blackwell

The “old friends” hypothesis suggests that exposure to a diverse range of co-evolved microbes during early life can regulate the immune system and reduce the risk of allergies. Previous research has shown that household factors are associated with the diversity and composition of the infant gastrointestinal microbiome (IGM). However, there is still limited data on different socio-environmental factors within household that may be associated with the development and composition of the IGM. The fecal microbiome (IFM) is used as an alternative to denote the gastrointestinal microbiome, a common tradition in microbiome literature. Here, we investigate potential relationships between household socio-environmental factors and infant fecal microbiome (IFM) composition and alpha diversity. We analyzed fecal samples from 48 healthy, exclusively breastfed infants from eastern Washington and northwest Idaho. The V1-V3 region of the bacterial 16S rRNA gene was sequenced to describe and analyze the role of household socio-environmental factors on the IFM. Household socio-environmental factors were reported by mothers of infants via survey and through naturalistic observations. Results showed that the household socio-environmental factors in our study were related to the composition of IFM but not the diversity of the IFM.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

Dedication

To my family, friends, mentors, and those who are curious.

CHAPTER ONE

INTRODUCTION

Socio-cultural and environmental factors in infancy are associated with health outcomes linked to the immune system. This is referred to as the "hygiene hypothesis" (Strachan, 1989). The "hygiene hypothesis" explained that in a multi-child household, exposure to infections in early childhood may leave children at a reduced risk of allergic diseases like asthma and hay fever later in their lives (Strachan, 1989).

Follow-ups to the original "hygiene hypothesis" have theorized that an infant's immune system can be affected by exposure to household factors (Alexandre-Silva et al., 2018). Larger families and more siblings maximize the chance of contact and transmission of infections from other people. In line with the hypothesis and the influence of infections, various infections can regulate the human immune system (Alexandre-Silva et al., 2018). The prediction about the influence of exposure to infections from environmental factors like household environment on immune system function, is especially useful when it comes to predicting allergy symptoms including hay fever, eczema, and asthma in childhood, although Strachan admitted that household size is not the only predictor for the prevalence of atopy (Strachan, 2000). Additionally, infants' early exposure to the diversity and the composition of specific species of microbes can influence gastrointestinal colonization (Alexandre-Silva et al., 2018; Debray et al., 2022). The early colonization by some bacteria can affect the further composition of other bacteria in the gastrointestinal environment (Debray et al., 2022).

The "old friends" hypothesis, an attempt to revise the "hygiene hypothesis" from an evolutionary perspective, suggested that exposure to a different range of co-evolving bacteria,

which may not always be infectious, can be helpful for immune system recognition and reducing the likelihood of allergies and autoimmune conditions (Rook et al., 2003; Rook, 2012). The "old friends" are bacteria that have co-evolved with humans for a long time, especially during the thousands of years of the Neolithic period when humans domesticated and lived with animals in agricultural settings (Rook, 2012). In support of the "old friends" hypothesis, studies have revealed that exposure to co-evolving and beneficial bacteria in the household may promote better immune function and lower the chances of allergies in children (Ownby et al., 2002). The preventive effect of co-evolving bacteria on reduced risk of allergic diseases may be due to the modulation of the infant's immune system. By repeated exposure to the co-evolved bacteria (old friends), these co-evolved bacteria have become part of human physiology and shaped the gastrointestinal microbiome to educate the immune system to not overreact against allergens (Rook et al., 2003).

Education of infants' immune systems can occur through exposure to microorganisms including bacteria, and the development of differentiation between harmful and harmless antigens (Alexandre-Silva et al., 2018). Thus, through this process, the immune system also learns to recognize and eliminate harmful microorganisms and tolerate the microbiome (as the commensal organisms within gastrointestinal tract) (Gensollen et al., 2016), and not overreact to allergens. For example, studies on germ-free animals that are not exposed to environmental factors and microorganisms have shown some impairments in their immune system and immune functions, including a decreased number of immune agents in the intestinal tract (Gensollen et al., 2016; Umesaki et al., 1993). These immune agents including CD8+ and CD4+ T cells, are responsible for detecting and responding to a variety of pathogens. Other impairments link to the

levels of αβ T cells and an uneven balance of T helper 2 (Th2) and T helper 1 (Th1) cells (Gensollen et al., 2016; Umesaki et al., 1993). Th2 cells are essential immune agents due to their role in protecting against parasites and immune responses linked to allergies and autoimmune conditions (Walker & McKenzie, 2018; Zheng et al., 2009).

Reducing the impact on the immune system and helping to educate the immune system in germ-free adult mice, can happen by reintroducing a standard microbiome from mice or humans (Gensollen et al., 2016). This reintroduced microbiome should consist of a diverse population of bacteria. Diversity is essential due to its role in regulating not only the number, and imbalances of T cells but also their responses to germ-free animals (Gaboriau-Routhiau et al., 2009; Gensollen et al., 2016). Regulatory T cells (Tregs) are the immune system's agents that enable immune system to tolerate non-self-antigens, including commensal bacteria and allergen antigens existing in the environment (Gensollen et al., 2016). Thus, one possible pathway to educate the human immune system to not overreact against allergens is to expose the immune system to a diversity of bacteria shaped by environmental exposure in early life. In addition, the early composition of gastrointestinal bacteria can affect the diversity of gastrointestinal bacteria (Debray et al., 2022). Hence, both the diversity and composition of the infant gastrointestinal microbiome (IGM) are crucial for the education of the immune system.

The diversity and composition of the gastrointestinal microbiome may differ among individuals due to the influence of different early socio-environmental exposures in infancy (Manus et al., 2023; Thompson et al., 2015; Yasmin et al., 2017). Moreover, other factors like age, genetic background, the geographical location people live, what they eat, and how often

they take antibiotics can have an impact on both the diversity and composition of the gastrointestinal microbiome (Benson et al., 2010; Yasmin et al., 2017; Yatsunenko et al., 2012).

The formation of the gastrointestinal microbiome in humans predominantly occurs from birth to three years of age; or not, the colonization might start from the fetal period (Lauder et al., 2016; Palmer et al., 2007; Yatsunenko et al., 2012). During the early years of life, the microbiome goes through stages of development. These stages are crucial as they enable the establishment of bacteria in the gastrointestinal tract, and they can also be the foundation for forming the microbiome and support the maturation of infants' immune systems (Stewart et al., 2018). Moreover, the fecal sample is often used for research on gastrointestinal microbiome. Thus, the " fecal microbiome" referred to as an alternative to the "gastrointestinal microbiome", which is a common tradition in the literature on gastrointestinal microbiomes (Amato, 2017).

The household sociocultural environment is among the first important sources of infants' early exposure to microorganisms shaping the diversity and composition of the infant fecal microbiome (IFM) (Azad et al., 2013; Lane et al., 2019). Consequently, household environment can affect infants' immune systems and the likelihood of developing allergies through effects on the gastrointestinal/ fecal microbiome. For example, the presence of pets in the household has been associated with increased gastrointestinal microbiome diversity and the increased relative abundance of taxa like *Bifidobacterium pseudopodium* in infants (Kim et al., 2019), *Oscillospira*, and *Ruminococcus*, which can be negatively linked to childhood atopy (Tun et al., 2017).

Nevertheless, when infants were exposed to different household microbial factors like pets and siblings, it was found that pets in the household minimized the abundance of *Bifidobacteriaceae* but increased the abundance of *Peptostreptococcaceae*; the diversity and

richness of IFM were high (Azad et al., 2013). On the contrary, in the same study, living with older siblings lowered the abundance of *Peptostreptococcaceae* in IFM and decreased the richness and diversity of IFM (Azad et al., 2013). Thus, co-residency with siblings and pets may be associated with the diversity and composition of the IFM.

Some research has studied the influence of household factors (aside from pets and siblings) like the size of the household (number of resident people in the house) and extended family on different alpha diversity indices (evenness, diversity, and richness) and composition of IFM. For example, Lane and colleagues (2019) found that the household composition (count of siblings and other members of the household) did not affect the alpha-diversity of IFM, but siblings' presence increased the abundance of *Lactobacillus*. Additionally, the presence of extended family in the household was related to the abundance of *Clostridium* sensu stricto 1 (Lane et al., 2019).

Moreover, research showed that social factors such as the number of non-maternal caregivers (allomothers) play a role in shaping the variety and makeup of bacterial communities found in the milk (Meehan et al., 2018) and IFM (Wiley et al., 2024). Studies indicated that having allomothers is associated with increased microbiome diversity in mother's milk and infant's skin while household size impacts the types of bacteria on the infant skin microbiome (Manus et al., 2023). However, how these factors influence the infant's gastrointestinal microbiome remains unclear (Wiley et al., 2024). It is suggested that caregiving practices and physical contact could directly impact the infant's microbiome through interaction with caregivers' skin and through the milk microbiome (Wiley et al., 2024). Thus, these findings demonstrated the importance of caregiving practices in colonizing the infant's gastrointestinal

microbiome. Thus, in light of the potential association between the early household socio-environmental factors and the development of infant's gastrointestinal microbiome, our objective is to address how socio-environmental factors in households might relate to 1) diversity and 2) composition of IFM. The early household socio-environmental factors in our study include physical interactions with caregivers, the number of different categories of non-household residents visiting the house since birth, the number of different categories of caregivers who care for the infant in mother's absence, presence of animals in the house (cat or dog, or others), co-sleeping with mother, household density (household size divided by bedrooms), and mothers working outside home.

CHAPTER TWO

RESEARCH DESIGN AND METHODOLOGY

Study populations

Our study's population consisted of 51 mother-infant pairs from rural areas of eastern Washington and northwest Idaho who participated in the Mother-Infant Microbiomes, Behavior, and Ecology Study (MIMBES). The infants were exclusively fed human milk, either via the breast or via bottle feeding. All infants were ≤ 6 months old and the range of their birth weight was 2.5-5 kilograms. The infants were reported by their mothers as healthy.

Metadata

Mothers provided demographic, household and caregiving practice information via questionnaires. Questions included mothers' age, ethnicity, and the delivery mode of the infant (c-section or vaginal). They were also asked about whether they worked outside the home(whether part-time, full-time, or volunteer), whether infants slept with their mother (co-sleeping), people who were not household residents had visited the house since the infant was born, people who cared for children in the absence of the mother, infants' age, keeping animals inside the house (including dogs, cats, or others), household size, and the number of bedrooms. Washington State University Institutional Review Board reviewed and approved the study (#15852).

Observational data

For the infants' observational data, focal observations were carried out on all participating infants. These observations were conducted by trained research assistants over three days from 7 am to 7 pm. During this time infants' physical interactions with their caregivers were recorded. The observations were separated into 4-hour segments each day. After 45 minutes of observation, the observer took a 15-minute break to prevent observer fatigue. The data documented 9 hours of infant behavior and interactions spread over 12 hours.

To identify and analyze the frequency of allomother (nonmaternal) physical contact, The frequency of allomother physical contact quantified any form of physical interaction, including holding or touching of the infant by individuals other than the mother in every 30-second interval of observation. Each person interacting with the infant was given an identifier to help study caregivers' interactions with the child, as described in the cohort study (Holdsworth et al., 2023). We grouped caregivers into 11 categories: partner/ father, grandmother, brother, sister, young girl, young boy, woman, man, older woman, and older man. The youth, adult and elderly groups included relatives (cousins, aunts, uncles) and non-family members. These categories were exclusive to ensure that each person belonged to one group (Holdsworth et al., 2023). The frequencies of physical contact, measured in seconds, were converted to minutes. Finally, the observational data was used to determine the frequency of allomother physical contact.

Household socio-environmental variables

After gathering the metadata and observational data, the household socio-environmental variables included the frequency of allomother physical contact with the infant (the only variable from observational data), presence of animals in the house (cat or dog, or others), infant co-

sleeping with mother, and whether mothers work outside the home (part-time, full-time, volunteer). The household density was defined as the number of household's residents divided by number of bedrooms (household size/bedrooms). In addition, the caregiver network (husband/partner, siblings, grandmother, grandfather, another family member, non-related caregiver, daycare provider, other) was defined as distinct categories of people identified by the mother as caregivers for her infant in her absence. Non-household visitors (siblings other than those who reside in the home, grandmother, grandfather, another family member, non-related caregiver (e.g., babysitter), friends, unrelated children, other) were defined as the number of different categories of non-household people who visited the house since the infant was born. Thus, we use the terms caregiver network for the number of different categories of people that the mother identified as caregivers for her infant in her absence, non-household visitors for the number of different categories of non-household people who visited the house since the infant was born, and household density for the household size divided by number of bedrooms.

Fecal samples collection

A fecal sample from each infant was collected in order to measure the infant fecal microbiome. A researcher removed the old diaper, wiped off the infant with a castile soap wipe, and put on the study-provided diaper. If the infant defecated in the diaper in the researcher's presence, the researcher wore new gloves and used the collection tube and spatula to scoop as much fecal material as possible into the tube while not exceeding half of the tube volume. If the infant did not defecate in the presence of the research assistant, they provided the mother with a fecal collection kit at home, including diapers, castile soap wipes, and gloves. The mother was shown how to replace a wet diaper and clean the newborn with a castile soap wipe as a last step

before changing to a fresh study diaper. When the infant had defecated, the mother put on gloves and accumulated as much fecal material as possible into the collecting tube with the spatula. Samples were placed in the home freezer and were picked up by research personnel and kept frozen at -20 °C until analyzed.

Infant fecal sample DNA extraction

The fecal samples were allowed to thaw on ice. After thawing, 200 mg of each sample was placed into a sterile microcentrifuge tube, and 0.5 mL of TE50 (10 mM Tris-HCl, 50 mM EDTA, pH 8) was added. Samples were vortexed to give an even combined, then stored at -80°C until further processing. DNA extraction was done using the QIAamp Fast DNA Stool Mini Kit (Qiagen, Germantown, MD), as previously described (Lackey et al., 2019). DNA was eluted with 200 µL of ATE buffer (supplied in kit) and stored at -80 °C until PCR amplification.

Bacterial DNA amplification

The process of amplifying DNA followed the method outlined in a study by Lackey and colleagues (Lackey et al., 2019). The amplification involved using primers to target and amplify the V1-V3 region of the 16S rRNA gene through a barcoded two-step, 30-cycle polymerase chain reaction (PCR). The PCR amplicons underwent cleaning and sizing using beads sourced from the HighPrep PCR Clean-up System by MagBio Genomics Inc. (MagBio Genomics Inc., Gaithersburg, MD) following the guidelines provided by the manufacturer. We checked cleaned amplicons utilizing the Accuclear Ultra High Sensitivity dsDNA Quantitation Kit from Biotium (Biotium, Fremont, California). The cleaned amplicons were pooled to ensure containing 25 ng of DNA per sample. These pools (amplicon pools) underwent another purification step with beads verified for quality using a Fragment Analyzer (Advanced Analytical Technologies,

Ankeny, IA, USA). Further, their quality was checked with the KAPA Biosciences Illumina library quantification kit and the Applied Biosystems StepOne Plus real time PCR equipment. The Genomics and Bioinformatics Resources Core, at the University of Idaho conducted sequencing on the combined samples (pools) using an Illumina MiSeq (San Diego, California). They employed v3 paired-end 300-bp protocol spanning 600 cycles.

Bacterial DNA sequencing

We demultiplexed the sequencing reads by dbcAmplicons (https://github.com/msettles/dbcAmplicons). We processed demultiplexed sequencing reads with DADA2 (v 1.16.0; Callahan ref) and decontam (v 1.8.0; Davis ref) packages in R (v 4.2.2; R Core Team, 2022) as described previously (Pace et al., 2021).

We utilized scaling sample read counts of each taxon by the total read counts in the sample to calculate the relative abundances of bacterial taxa at genus level. Before calculating diversity indices and evaluating rarefaction curves conducted at the amplicon sequence variant (ASV) level, sample read counts were rarefied to a sampling depth of 773 reads, removing three samples of reads < 773 from the dataset. As a result, the analytic sample consisted of 48 infants.

Statistical analysis

For the calculation of observed richness and Shannon diversity (two indices of alpha diversity measures), we used the estimate_richness function in the vegan package in R (v 2.5-7; Oksanen ref). The formula found in (Lin & Peddada, 2024) the book "Numerical Ecology with R" was used to calculate Pielou's J evenness.

We conducted other statistical analyses using R version 4.3.1.; R Core Team (2024). After reviewing the histogram and summary statistics of the variables, we applied the winsorization on the frequency of allomother physical contact and the household density. We winsorized the frequency of allomother physical contact at 5% and 95% of the distribution to reduce the impact of outliers. The cut-offs were < 0.50 and >106.02, resulting in winsorizing five samples from 48 samples. The household density was also winsorized at 5% and 95% of the distribution. The cut-offs for the household density were < 1.0 and >2.5, which led to the winsorization of the three samples from 48 samples.

Regression models

For each of the three dependent variables (observed richness, Shannon diversity, and Pielou's J evenness), we ran seven separate regression models with each household socio-environmental variable while controlling for delivery mode and infant sex. In total, we ran 21 regression models. We used the Benjamini-Hochberg method to adjust p-values and restrain the effect of the false discovery rate (FDR), which is used in multiple-hypothesis testing. Significant thresholds were set at the adjusted p-value <0.05.

Differential abundance of bacterial genera in the fecal microbiome

We used the function of ANCOMBC2 from ANCOM-BC (Analysis of Compositions of Microbiomes with Bias Correction) package version 7.4.0 in R version 4.3.1 (Lin & Peddada, 2024) to assess if the abundance of genera in the IFM differed based on the household ecology factors. The function of ANCOM-BC2 analyzes the differential abundance of taxa while accounting for not only sample-specific bias and false discovery error but also for taxon-specific bias common in microbiome studies. The unrarefied read count data of samples with > 773 reads

were used in ANCOM-BC2. It allowed us to create regression models predicting the absolute abundance differences of each genus (differential abundance) with each independent household socio-environmental variable separately while controlling for infant sex and delivery mode. The models for each household socio-environmental variable were corrected for multiple testing by the Holm–Bonferroni method. Significance was determined as adjusted $p<0.05$ for all statistical tests conducted by ANCOM-BC2. The taxa with a high proportion of zeros (more than 0.90) were not included in the analysis, and all other inputs were the defaults from ANCOM-BC2.

CHAPTER THREE

RESULTS

Study population summary

The mean age of mothers was 29 years old (mean ± s.d.;29.2 ± 3.6). Mothers who worked outside the home made up 48% of the population in our study. The eighty two percent (n=39) of mothers identified themselves as Caucasian/European American, 8% (n=4) Asian/Asian American, 4% (n=2) Hispanic/Latino, and 6% (n=3) identified themselves as other ethnic groups (Table 1).

Fifty percent of infants were female, and 50% were male (total n=48). The range of infants' age was between 28 days to 6 months (10.7 ± 5.9 weeks, and the mean weight of infants at birth was 3.5 kilograms (3.5 ± 0.4), and 73% of infants were born vaginally. The range of household density was 1 to 4 people per bedroom (1.6 ± 0.6). The number of different categories of non-household visitors after the infant was born ranged from 1 to 6 (4.2 ± 1.2). The range for the number of different categories of people who care for the infant in the mother's absence was 0 to 3 (1 ± 1). Forty-four percent of households had animals (mostly cats and/or dogs) in their household (Table 1).

Descriptive summary of diversity indices and taxa composition

Among the 48 samples with read counts > 773, the mean richness of the samples was 34 ± 11.8 the Shannon diversity of samples was 2.3 ± 0.4**,** and Pielou's J evenness of samples 0.7 ± 0.1 (Table 1).

Table 1: Descriptive statistics of study population (N=48), and alpha diversity indices.

Variable	N (%)	Variable	Mean (SD)	Range
Mother Working Outside Home /Yes	23 (48 %)	Household Density	2 (0.6)	1- 2.5
Animals in House /Yes	21 (44 %)	Caregivers Network	1 (1)	0-3
Cosleeping With Mother/Yes	24 (50 %)	Mom Age (Year)	29 (4)	23- 38
Infant Sex /Male	24 (50 %)	Infant age (Weeks)	11 (6)	4- 25
Infant Sex /Female	24 (50 %)	Frequency of allomother physical contact	30 (31)	0.5- 113
Mom Ethnicity (Hispanic/Latino)	2 (4 %)	Non-household Visitors	4 (1)	1- 6
Mom Ethnicity (Caucasian/European American)	39 (81 %)	Infant birth weight (kg)	3 (0.4)	2- 5
Mom Ethnicity (Asian/Asian American)	4 (8 %)	Observed richness	33.98 ± 11.78	
Mom Ethnicity (Others)	3 (6 %)	Shannon diversity	2.28 ± 0.42	
Delivery Mode/ Cesarean	13 (27 %)	Pielou’s J evenness	0.66 ± 0.10	
Delivery Mode/ Vaginal	35 (73 %)			

Bifidobacterium had the highest mean ± s.d. (19 ± 25%) of taxonomic relative abundance in all the samples. After that, there were *Veillonella* **(**16 ± 24%), *Bacteroides* (15 ± 22%), *Escherichia/Shigella* (11 ± 19%), *Clostridium sensu stricto 1* (6 ± 15%), *Ruminococcus gnavus group* (5 ± 15%), *Klebsiella* (5 ± 13%), *Akkermansia* (3 ± 12%), *Citrobacter* (3 ± 10%), *Enterobacter* (2 ± 6%), *Streptococcus* (2 ± 5%), *Erysipelotrichaceae UCG-003* (1 ± 9%), *Erysipelatoclostridium* (1 ± 6%), *Alistipes* (1 ± 6%), *Pluralibacter* (1 ± 6%), unspecified genera in the family *Enterobacteriaceae* (1 ± 4%), and *Parabacteroides* (1 ± 3%). The relative abundance of other genera was less than <1%. The relative abundance of each genus in each sample is depicted in Figure 1.

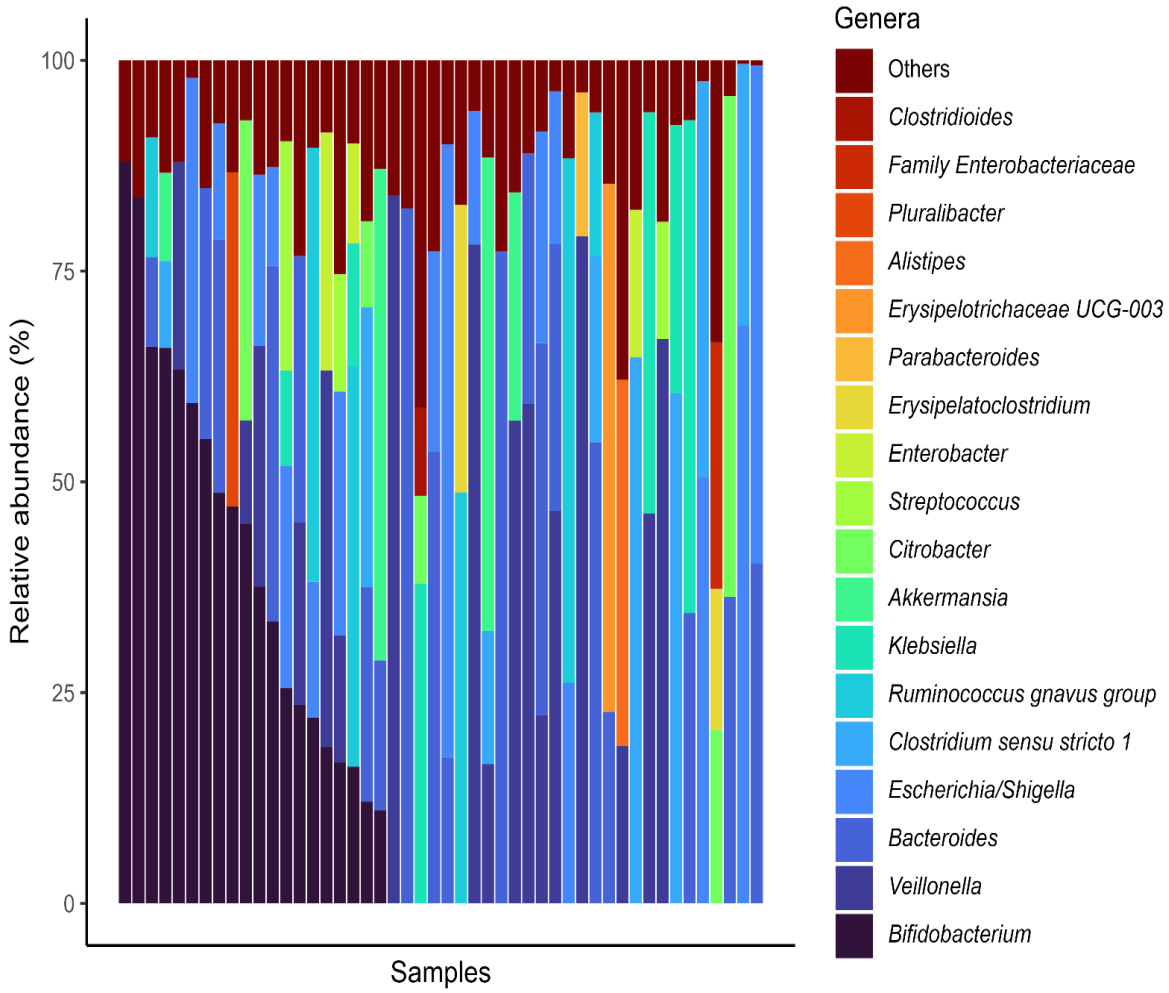

Figure 1: Relative abundance of genera in each infant fecal sample (N=48). Samples are sorted from highest to lowest relative abundance of *Bifidobacterium*. The "Others" are genera with a relative abundance <1%.

Household socio-environmental variables and alpha diversity of IFM

Regression models assessed potential association between the socio-environmental variables in household and observed richness, Shannon diversity, and Pielou's J evenness while controlling for infant sex and delivery mode. After p-value adjustment, we found that there were no statistically significant relationships between any household socio-environmental variables and any alpha- diversity indices (Table 2a, 2b, and 2c)

Table 2a: Linear regression models examining relationship of household socio-environmental variables with observed richness.

Observed richness			
Variable	Coefficient (Standard error)	Adjusted P-value	Adjusted R^2
Frequency of allomother physical contact	-0.002 (0.06)	0.97	-0.03
Household density	-0.04 (3.98)	0.99	-0.03
Co-sleeping with mother	-1.87 (3.48)	0.70	-0.02
Mother working outside home	-0.33 (3.55)	0.93	-0.03
Non-household visitors	-3.20 (1.40)	0.08	0.08
Caregiver network	1.61 (2.00)	0.62	-0.02
Animals in house	-3.30 (3.45)	0.52	-0.01
* All models controlled for infant sex and delivery mode.			

Table 2b: Linear regression models examining relationship of household socio-environmental variables with Shannon diversity.

Shannon diversity			
Variable	Coefficient (Standard error)	Adjusted/P-value	Adjusted R^2
Frequency of allomother physical contact	-0.00 (0.00)	0.69	0.05
Household density	0.00 (0.13)	0.99	0.05
Co-sleeping with mother	-0.07 (0.12)	0.62	0.06
Mother working outside home	0.06 (0.12)	0.75	0.05
Non-household visitors	-0.03 (0.05)	0.61	0.06
Caregiver network	0.09 (0.07)	0.29	0.08
Animals in house	-0.00 (0.12)	0.99	0.05
* All models controlled for infant sex and delivery mode.			

Table 2c: Linear regression models examining relationship of household socio-environmental variables with Pielou's J evenness.

Pielou's J evenness			
Variable	Coefficient/ Standard error	Adjusted/P-value	Adjusted R^2
Frequency of allomother physical contact	-0.00 (0.00)	0.54	0.15
Household density	0.00 (0.03)	0.79	0.15
Co-sleeping with mother	-0.01 (0.26)	0.81	0.15
Mother working outside home	0.02 (0.03)	0.42	0.16
Non-household visitors	0.01 (0.01)	0.53	0.15
Caregiver network	0.01 (0.01)	0.36	0.16
Animals in house	0.01 (0.03)	0.57	0.15
* All models controlled for infant sex and delivery mode.			

Household socio-environmental variables and differential abundance of bacterial genera in IFM

ANCOM-BC2 analyses showed that the differential abundance of some of the 44 genera were significantly associated with household socio-environmental variables at $q<0.05$ (adjusted p-value) (Figure 2). The log fold change (LFC) was utilized to demonstrate the magnitude and direction of these associations (Figure 2). Two genera out of 44 were differentially abundant by household density. These genera included *Akkermansia* (LFC= -9.79, q= 0.00), *Parabacteroides* (LFC= 1.91, q= 0.02)

The genera *Citrobacter* (LFC=2.62, q= 0.008) and *Enterobacter* (LFC= 2.08, q= 0.002), were estimated to have increased abundance in infants whose mothers worked outside the home, while *Lactobacillus* (LFC= -3.48, q= 0.00), *Actinomyces* (LFC= -1.9, q= 0.02), and *Atopobium* (LFC= -2.64, q= 0.03) were estimated to have decreased abundance among infants whose mothers worked outside the home

The frequency of allomother physical contact was significantly associated with the abundance of 7 genera, including positive association with *Lachnoclostridium* (LFC= 0.06 q= 0.0003), *Sutterella* (LFC= 0.06, q= 0.004), *Atopobium* (LFC= 0.08, q=0.001), *Akkermansia* (LFC= 0.04, q= 0.02), and *Faecalibacterium* (LFC= 0.08, q= 0.001); negative association with *Lactobacillus* (LFC= -0.07, q= 0.00), *Actinomyces* (LFC= -0.03, q= 0.03),.

The presence of the animals in the house (cat or dog, or others) was associated with the increased abundance of *Ruminococcus* gnavus group (LFC= 3.06, q= 0.02), and decreased abundance of *Eggerthella* (LFC= -1.7, q= 0.04).

Co-sleeping with mother was associated with decreased abundance of *Lactobacillus* (LFC= -3.53, q= 0.0004), and *Akkermansia* (LFC= -4.66, q= 0.0008), and increased abundance

of *Faecalibacterium* (LFC= 3.57, q= 0.01), *Kosakonia* (LFC= 2.35, q= 0.0007), *Haemophilus* (LFC= 3.64, q= 0.0005), *Atopobium* (LFC= 2.02, q= 0.04), and *Sutterella* (LFC= 2.22, q= 0.03).

The caregiver network was, positively, related to the differential abundance of *Hungatella* (LFC= 2.13, q= 0.000), *Lacticaseibacillus* (LFC= 0.74, q= 0.0008), *Actinomyces* (LFC= 1.71, q= 0.0001), *Rothia* (LFC= 0.81, q= 0.0005), *Atopobium* (LFC= 0.88, q= 0.02), *Akkermansia* (LFC= 1.12, q= 0.03), and *Flavonifractor* (LFC= 2.16, q= 0.02); and negatively, related to the differential abundance of *Ruminococcus* gnavus group (LFC= -1.23, q= 0.01), *Megasphaera* (LFC= -2.10, q= 0.003), *Lachnoclostridium* (LFC= -2.37, q= 0.000), *Clostridioides* (LFC= -3.65, q= 0.0009), *Corynebacterium* (LFC= -1.13, q= 0.01), *Eggerthella* (LFC= -0.85, q= 0.0005), and *Faecalibacterium* (LFC= -1.49, q= 0.000).

The non-household visitors had negative relationship with the differential abundance of *Lachnoclostridium* (LFC= -1.64, q= 0.03), *Hungatella* (LFC= -1.3, q= 0.03), *Ruminococcus* gnavus group (LFC= -2.05, q= 0.03), and positive relationship with the differential abundance of *Akkermansia* (LFC= 1.7, q= 0.03).

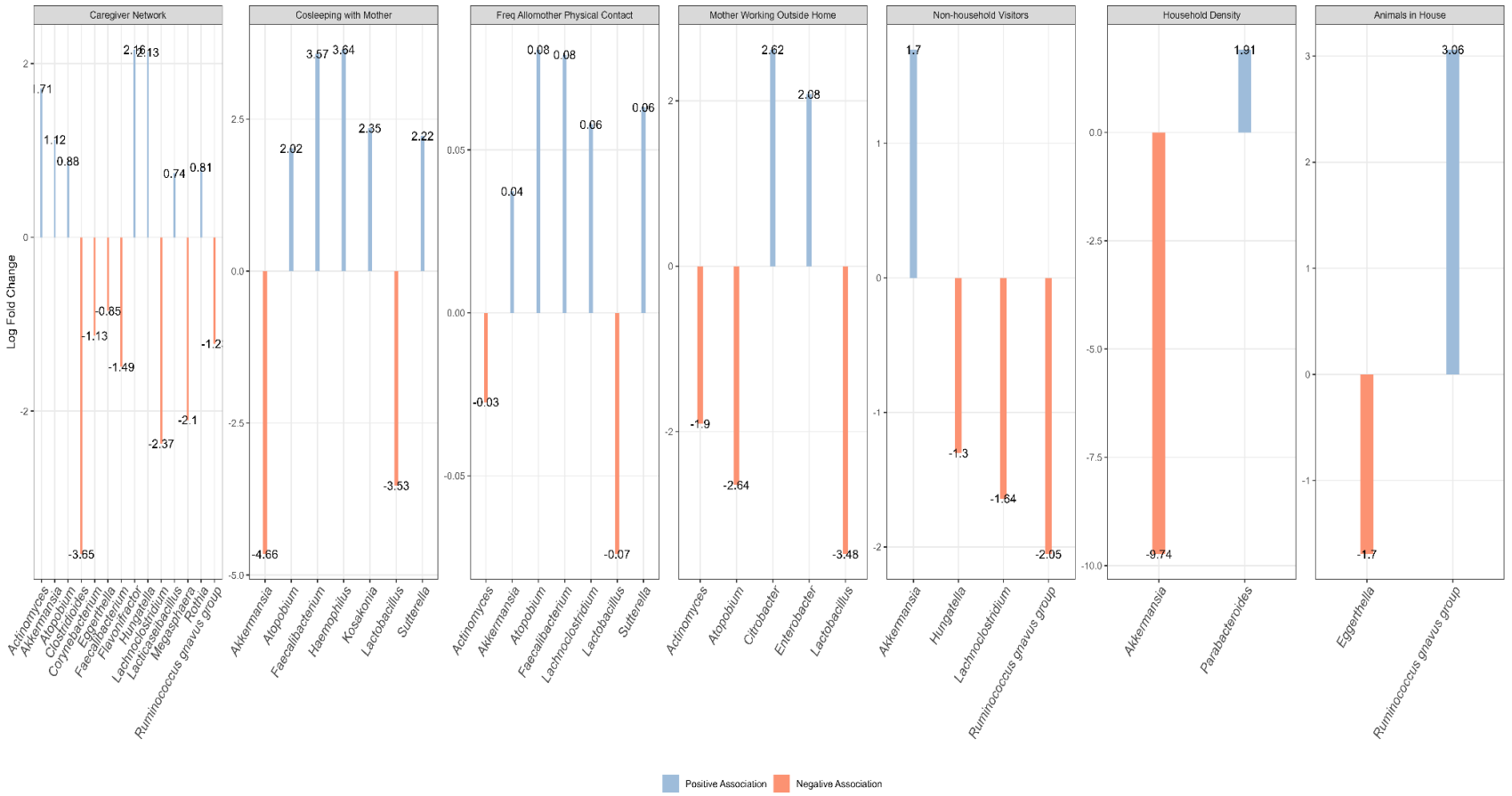

Figure 2: Log fold change estimates the direction and magnitude of the association between the abundance of significant genera with each socio-environmental variable in household ($q < 0.05$).

CHAPTER FOUR

DISCUSSION

In this study, we found that 1) Socio-environmental factors within the household were not associated with any of the three alpha diversity indices of the infant fecal microbiome when controlled for infant sex and delivery mode. The socio-environmental factors within the household in our study included physical interactions with caregivers, the number of different categories of non-household residents visiting the house since the infant's birth, the number of different categories of caregivers who care for the infant in the absence of the mother, the presence of animals in house (mostly cat and/or dog), co-sleeping with mother, household density (household size divided by bedrooms), and mothers working outside home. 2) On the other hand, all socio-environmental factors within the household were associated with differential abundance of mostly genera with low relative abundance, also controlling for infant sex and delivery mode.

IFM characteristics compared to other populations

Regarding the IFM description in our study, *Bifidobacterium, Veillonella, Bacteroides*, *Escherichia/Shigella*, and *Clostridium* sensu stricto 1 had the highest relative abundance, similar to Lane and colleagues' study that found these genera were abundant in all their study populations (Lane et al., 2019). Similarly, the relative abundance of *Bifidobacterium* was high in Azad and colleagues' study (Azad et al., 2013). The relative abundance of *Lactobacillus*, one of the abundant genera in most studies, was less than 1% in our study. Similarly, Lackey and colleagues revealed that the relative abundance of *Lactobacillus* in populations in Sweden,

Spain, the United States, and Peru was low, while in the rural populations of the Gambia and Ethiopia, it was the highest (Lackey et al., 2019).

In terms of comparing our result for diversity with the other studies, our study revealed that the mean for Shannon diversity was 2.28, and observed richness was 33.98. The results of other studies regarding Shannon diversity and richness were relatively similar to our study (Azad et al., 2013; Lackey et al., 2019). The mean for Shannon diversity in Azad and colleagues' study was 1.4 when they used QIIME to generate the Shannon diversity index with rarefied data (10,000 sequences per sample) in 24 samples (Azad et al., 2013). The mean of Shannon diversity and richness were 2.37 and 34, respectively, in the US/Washington for Lackey and colleagues' study (Lackey et al., 2019). They excluded the reads counts lower than 270 and used phyloseq to calculate diversity indices. To put it briefly, although there is variation in diversity and microbial communities in infant feces in different populations, the diversity and microbial communities found in the feces samples gathered for our research were relatively similar to previous studies.

Household socio-environmental variables and diversity and composition of IFM

Our models showed no associations between the household socio-environmental variables and alpha diversity. However, the models for the differential abundance of specific genera in the microbiome indicated the association between some genera's abundance and socio-environmental factors in the household.

Among all the household socio-environmental variables, the number of different categories of caregivers (caregiver network) was associated with the differential abundance of 14 genera, while the household density and the presence of animals (mostly cats and/or dogs) in the house were associated with the differential abundance of only two genera. The differential

abundance of *Akkermansia* was associated with all the household socio-environmental variables except for the presence of animals in the house and mothers working outside the home. The direction of the relationship was not the same for different household socio-environmental factors. It was positive for the frequency of allomother physical contact, caregiver network, and non-household visitors but negative for co-sleeping with mother and household density.

The caregiver network and frequency of allomother physical contact were not associated with the alpha diversity in our study. Nonetheless, the caregiver network was positively associated with the differential abundance of 7 genera including *Akkermansia* (LFC= 1.2), *Atopobium* (LFC= 0.88), *Flavonifractor* (LFC= 2.16), and *Hungatella* (LFC= 2.13), and negatively with differential abundance of 7 genera including *Clostridioides* (LFC= -3.65), *Faecalibacterium* (LFC= -1.49), and *Lachnoclostridium* (LFC= -2.37). Some strains of *Clostridioides* are linked to causing issues like diarrhea (Hernandez et al., 2019).

In addition, the frequency of allomother physical contact was positively associated with differential abundance of 5 genera including *Akkermansia* (LFC= 0.04), *Atopobium* (LFC= 0.08), *Faecalibacterium* (LFC= 0.08), *Sutterella* (LFC= 0.06), and *Lachnoclostridium* (LFC= 0.06). In composition, it was negatively associated with differential abundance of 2 genera including *Lactobacillus* (LFC= -0.07) and *Actinomyces* (LFC= -0.03).

Other research has studied caregiving practices like physical and skin-to-skin contact, and co-sleeping with caregivers (Manus et al., 2023b.; Wiley et al., 2024), due to the importance of caregiving in the colonization of IFM (Wiley et al., 2024). Wiley and colleagues studied the influence of physical contact with caregivers on the alpha (within individuals) and beta (between individuals) diversity of IFM. They used rarefied data lower 7295 read count for calculating the

alpha diversity (Shannon diversity and Chao 1 richness) with QIME 1.9.1. Contrary to our results, Wiley and colleagues found a relationship between the alpha diversity of IFM at six months of age and the average time related to physical contact with caregivers in early life (Wiley et al., 2024). However, they did not report the direction of the relationship between alpha diversity of IFM at six months of age and average time related to physical contact with caregivers in early life.

Co-sleeping with the mother, like frequency of allomother physical contact and caregiver network, was associated with the differential abundance of various genera, but the other variables were associated with the differential abundance of fewer genera. Co-sleeping with mother was positively associated with five genera, including *Atopobium* (LFC= 2.02), *Faecalibacterium* (LFC= 3.57), *Sutterella* (LFC= 2.22), and negatively with two genera, including *Lactobacillus* (LFC= -3.53) and *Akkermansia* (LFC= -4.66). Mothers working outside home were linked positively with the differential abundance of two genera, including *Citrobacter* (LFC= 2.62) but negatively with the differential abundance of the three genera *Lactobacillus* (LFC= -3.48), *Actinomyces* (LFC= -1.9), and *Atopobium* (LFC= -2.64). Consequently, it seems that, in our study, the influence of co-sleeping with mother on the differential abundance of taxa of IFM had similarities with the influence of caregiver network and frequency of allomother physical contact, and the influence of mothers working outside home seemed partially similar to the influence of the frequency of allomother physical contact on the differential abundance of similar taxa, regardless of their directions. Thus, some significant genera were shared among different household socio-environmental variables.

We found no link between the presence of animals (mostly cats and/or dogs) in the house and the alpha diversity of IFM, but animals in the house was related to the differential abundance of *Ruminococcus* positively (LFC= 3.06), and *Eggerthella* negatively (LFC= -1.7). In the context of comparing our results for the diversity of IFM, our result was inconsistent with Azad and colleagues' study (Azad et al., 2013). They checked the influence of pets in household on IFM and used QIME for calculating Shannon diversity with the data rarefied in 10,000 sequences per individual sample. They found that the presence of pets increases the diversity of IFM. In terms of comparing our result for influence of presence of animals in house on the composition of IFM and consistent with our study, Tune and colleagues showed that the presence of pets increased the abundance of *Ruminococcus,* although their methodology (multiple variable logistic regression) was different (Tune et al., 2017). *Ruminococcus* has been associated with minimizing the chance of atopy (Tune et al., 2017).

Non-household visitors (number of different categories of non-household residents who visited the house since infant's birth) positively linked to differential abundance of just one genus (*Akkermansia;* LFC= -2.37) and negatively to differential abundance of the three genera including *Lachnoclostridium* (LFC= -1.64).

We found no association between household density (household size/bedrooms) and alpha diversity. Our result for the association of household density and alpha diversity was partially consistent with the result from Lane and colleagues' study on populations from different countries, revealing no association between household composition (number of household residents including siblings and extended family) and alpha diversity in all study populations (Lane et al.,2019). Additionally, Wiley and colleagues found no significant association between

household size and alpha diversity of IFM (Wiley et al., 2023). Conversely, Manus and colleagues revealed that the number of people in a household during pregnancy can affect the alpha diversity of IFM (Manus et al., 2023). Interestingly, Manus and colleagues noted a negative relationship at two weeks old but a positive one, at six months of age.

Lane and colleagues discussed that co-residency, existing in a household, may influence the composition of IFM significantly more than IFM diversity and richness (Lane et al., 2019). Thus, in our study, co-residency existing in household socio-environmental factors may affect the composition of IFM more than IFM diversity and richness. The co-residency in our socio-environmental factors existed in household density (household size/bedrooms), co-sleeping with mother, the presence of animals in the house. Additionally, small household size in Western households, differences in hygiene practices of caring for infants, gender norms, and the influence of some bacteria on further colonization of the gastrointestinal environment may be a reason behind the limited support of our study from the association of socio-environmental household variables and alpha diversity (Debray et al., 2022; Jost et al., 2012; Lane et al., 2019; Wiley et al., 2024).

Out of the two genera of bacteria whose differential abundances were associated with household density, the differential abundance of one genus (*Akkermansia*) was negatively related to the household density with the log fold change of -9.79. Thus, in terms of the magnitude of the impact on the composition of IFM and compared to the other household variables, household density was assumed to have a considerable negative effect on the abundance of a specific genus (*Akkermansia*), which was highly abundant in IFM.

Additionally, in assessing the influence of household factors on the composition of IFM, it is essential to consider that the differential abundance of how many bacteria are significantly impacted by each factor and how many of the significant bacteria have an abundance of more than 1% in IFM. For example, among the household factors, caregiver networks affected a total of 14 bacteria, three of which are highly abundant in IFM, or household density was related to the differential abundance of two genera, all of which were abundant in IFM. Conversely, the frequency of allomother physical contact influenced seven genera. However, the majority of these genera are not abundant in IFM, suggesting that the overall effect of the frequency of allomother physical contact on microbiome composition may be more nuanced compared to factors associated with abundant genera, such as caregiver networks and household density.

As a result, regarding to A) the magnitude of influence of a household variable on differential abundance of a genus, B) the total number of the genus that their differential abundance is related to each variable, and C) the number of highly abundant bacteria that a household variable is linked to, it may be reasonable to conclude that caregiver networks and household density had a greater impact on composition of IFM compared to the other variables. However, we acknowledge that the relationship between the different bacteria in gastrointestinal tract is complicated, and there are numerous other factors that can affect the composition of IFM.

One of the limitations of our study was relying solely on 16S rRNA V3-V4 sequencing. 16S rRNA V3-V4 sequencing may not fully capture all the aspects of microbial diversity and composition that are crucial for a deep understanding of the microbiome (Hamady & Knight, 2009; Matchado et al., 2024). Some of the limitation stems from the inability of 16S rRNA V3-V4 gene sequencing, to provide detailed taxonomic identification down to the species and strain

level. Therefore, there is a need for approaches like shotgun metagenomics in future studies, that can offer precision and reveal compositional connections at both species and strain levels (Jovel et al., 2016). Additionally, the household socio-environmental factors in our study are not generalizable to all cultures and populations. We also posit that there are limitations in comparing our study with other research because there were differences across studies, based on their genetic differences, ethnicity, geographical regions, and infants' diet (exclusively or not exclusively breastfed). Moreover, there were differences between the methodologies. As outlined by Sharon and colleagues, it is worth mentioning that there is a need for quality control measures and standardized protocols between different studies to provide reproducible and comparable data on microbiome studies (Sharon et al., 2022). A strength of our study was related to using not only self-reported socio-environmental factors within the household but also using observational data assessing the frequency of physical contact in caregiving practices. The observations helped us with more understanding of caregiving characteristics.

In conclusion, the household socio-environmental factors within the household in our study were associated with the composition of the IFM but not the diversity of the IFM. Whether these differences in composition are functionally important for the development of the immune system and prevalence of disease remains to be investigated in future studies. However, the presence of differences suggests that the IFM remains a possible pathway for the influence of "old friends" on health.

REFERENCES

Alexandre-Silva, G. M., Brito-Souza, P. A., Oliveira, A. C. S., Cerni, F. A., Zottich, U., & Pucca, M. B. (2018). The hygiene hypothesis at a glance: Early exposures, immune mechanism and novel therapies. *Acta Tropica*, *188*, 16–26. https://doi.org/10.1016/j.actatropica.2018.08.032

Amato, K. R. (2017). An introduction to microbiome analysis for human biology applications. *American Journal of Human Biology*, *29*(1), e22931. https://doi.org/10.1002/ajhb.22931

Azad, M. B., Konya, T., Maughan, H., Guttman, D. S., Field, C. J., Sears, M. R., Becker, A. B., Scott, J. A., Kozyrskyj, A. L., & CHILD Study Investigators. (2013). Infant gut microbiota and the hygiene hypothesis of allergic disease: Impact of household pets and siblings on microbiota composition and diversity. *Allergy, Asthma & Clinical Immunology*, *9*(1), 15. https://doi.org/10.1186/1710-1492-9-15

Benson, A. K., Kelly, S. A., Legge, R., Ma, F., Low, S. J., Kim, J., Zhang, M., Oh, P. L., Nehrenberg, D., Hua, K., Kachman, S. D., Moriyama, E. N., Walter, J., Peterson, D. A., & Pomp, D. (2010). Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. *Proceedings of the National Academy of Sciences*, *107*(44), 18933–18938. https://doi.org/10.1073/pnas.1007028107

Debray, R., Herbert, R. A., Jaffe, A. L., Crits-Christoph, A., Power, M. E., & Koskella, B. (2022). Priority effects in microbiome assembly. *Nature Reviews Microbiology*, *20*(2), Article 2. https://doi.org/10.1038/s41579-021-00604-w

Gaboriau-Routhiau, V., Rakotobe, S., Lécuyer, E., Mulder, I., Lan, A., Bridonneau, C., Rochet, V., Pisi, A., De Paepe, M., Brandi, G., Eberl, G., Snel, J., Kelly, D., & Cerf-Bensussan, N. (2009). The Key Role of Segmented Filamentous Bacteria in the Coordinated Maturation of Gut Helper T Cell Responses. *Immunity*, *31*(4), 677–689. https://doi.org/10.1016/j.immuni.2009.08.020

Gensollen, T., Iyer, S. S., Kasper, D. L., & Blumberg, R. S. (2016). How colonization by microbiota in early life shapes the immune system. *Science*, *352*(6285), 539–544. https://doi.org/10.1126/science.aad9378

Hamady, M., & Knight, R. (2009). Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Research*, *19*(7), 1141–1152. https://doi.org/10.1101/gr.085464.108

Holdsworth, E. A., Williams, J. E., Pace, R. M., Lane, A. A., Gartstein, M., McGuire, M. A., McGuire, M. K., & Meehan, C. L. (2023). Breastfeeding patterns are associated with human milk microbiome composition: The Mother-Infant Microbiomes, Behavior, and Ecology Study (MIMBES). *PLOS ONE*, *18*(8), e0287839. https://doi.org/10.1371/journal.pone.0287839

Jost, T., Lacroix, C., Braegger, C. P., & Chassard, C. (2012). New Insights in Gut Microbiota Establishment in Healthy Breast Fed Neonates. *PLOS ONE*, *7*(8), e44595. https://doi.org/10.1371/journal.pone.0044595

Jovel, J., Patterson, J., Wang, W., Hotte, N., O'Keefe, S., Mitchel, T., Perry, T., Kao, D., Mason, A. L., Madsen, K. L., & Wong, G. K.-S. (2016). Characterization of the Gut Microbiome Using 16S or Shotgun Metagenomics. *Frontiers in Microbiology*, *7*. https://doi.org/10.3389/fmicb.2016.00459

Kim, H., Sitarik, A. R., Woodcroft, K., Johnson, C. C., & Zoratti, E. (2019). Birth Mode, Breastfeeding, Pet Exposure, and Antibiotic Use: Associations With the Gut Microbiome and Sensitization in Children. *Current Allergy and Asthma Reports*, *19*(4), 22. https://doi.org/10.1007/s11882-019-0851-9

Lackey, K. A., Williams, J. E., Meehan, C. L., Zachek, J. A., Benda, E. D., Price, W. J., Foster, J. A., Sellen, D. W., Kamau-Mbuthia, E. W., Kamundia, E. W., Mbugua, S., Moore, S. E., Prentice, A. M., K., D. G., Kvist, L. J., Otoo, G. E., García-Carral, C., Jiménez, E., Ruiz, L., … McGuire, M. K. (2019). What's Normal? Microbiomes in Human Milk and Infant Feces Are Related to Each Other but Vary Geographically: The INSPIRE Study. *Frontiers in Nutrition*, *6*. https://www.frontiersin.org/articles/10.3389/fnut.2019.00045

Lane, A. A., McGuire, M. K., McGuire, M. A., Williams, J. E., Lackey, K. A., Hagen, E. H., Kaul, A., Gindola, D., Gebeyehu, D., Flores, K. E., Foster, J. A., Sellen, D. W., Kamau-Mbuthia, E. W., Kamundia, E. W., Mbugua, S., Moore, S. E., Prentice, A. M., Kvist, L. J., Otoo, G. E., … Meehan, C. L. (2019). Household composition and the infant fecal microbiome: The INSPIRE study. *American Journal of Physical Anthropology*, *169*(3), 526–539. https://doi.org/10.1002/ajpa.23843

Lauder, A. P., Roche, A. M., Sherrill-Mix, S., Bailey, A., Laughlin, A. L., Bittinger, K., Leite, R., Elovitz, M. A., Parry, S., & Bushman, F. D. (2016). Comparison of placenta samples with contamination controls does not provide evidence for a distinct placenta microbiota. *Microbiome*, *4*(1), 29. https://doi.org/10.1186/s40168-016-0172-3

Lin, H., & Peddada, S. D. (2024). Multigroup analysis of compositions of microbiomes with covariate adjustments and repeated measures. *Nature Methods*, *21*(1), Article 1. https://doi.org/10.1038/s41592-023-02092-7

Manus, M. B., Sardaro, M. L. S., Dada, O., Davis, M. I., Romoff, M. R., Torello, S. G., Ubadigbo, E., Wu, R. C., Miller, E. S., & Amato, K. R. (2023b). Interactions with alloparents are associated with the diversity of infant skin and fecal bacterial communities in Chicago, United States. *American Journal of Human Biology*, *n/a*(n/a), e23972. https://doi.org/10.1002/ajhb.23972

Manus, M. B., Watson, E., Kuthyar, S., Carba, D., Belarmino, N. M., McDade, T. W., Kuzawa, C. W., & Amato, K. R. (2023). Prenatal household size and composition are associated with

infant fecal bacterial diversity in Cebu, Philippines. *American Journal of Biological Anthropology*, *181*(1), 45–58. https://doi.org/10.1002/ajpa.24720

Matchado, M. S., Rühlemann, M., Reitmeier, S., Kacprowski, T., Frost, F., Haller, D., Baumbach, J., & List, M. (2024). On the limits of 16S rRNA gene-based metagenome prediction and functional profiling. *Microbial Genomics*, *10*(2). https://doi.org/10.1099/mgen.0.001203

Meehan, C. L., Lackey, K. A., Hagen, E. H., Williams, J. E., Roulette, J., Helfrecht, C., McGuire, M. A., & McGuire, M. K. (2018). Social networks, cooperative breeding, and the human milk microbiome. *American Journal of Human Biology*, *30*(4), e23131. https://doi.org/10.1002/ajhb.23131

Ownby, D. R., Johnson, C. C., & Peterson, E. L. (2002). Exposure to Dogs and Cats in the First Year of Life and Risk of Allergic Sensitization at 6 to 7 Years of Age. *JAMA*, *288*(8), 963–972. https://doi.org/10.1001/jama.288.8.963

Pace, R. M., Chu, D. M., Prince, A. L., Ma, J., Seferovic, M. D., & Aagaard, K. M. (2021). Complex species and strain ecology of the vaginal microbiome from pregnancy to postpartum and association with preterm birth. *Med*, *2*(9), 1027-1049.e7. https://doi.org/10.1016/j.medj.2021.06.001

Palmer, C., Bik, E. M., DiGiulio, D. B., Relman, D. A., & Brown, P. O. (2007). Development of the Human Infant Intestinal Microbiota. *PLOS Biology*, *5*(7), e177. https://doi.org/10.1371/journal.pbio.0050177

Rook, G. (2012). A Darwinian View of the Hygiene or "Old Friends" Hypothesis. *Microbe Magazine*, *7*, 173–180. https://doi.org/10.1128/microbe.7.173.1

Rook, G. A. W. (2012). Hygiene Hypothesis and Autoimmune Diseases. *Clinical Reviews in Allergy & Immunology*, *42*(1), 5–15. https://doi.org/10.1007/s12016-011-8285-8

Rook, G. A. W., Martinelli, R., & Brunet, L. R. (2003). Innate immune responses to mycobacteria and the downregulation of atopic responses. *Current Opinion in Allergy and Clinical Immunology*, *3*(5), 337.

Sharon, I., Quijada, N. M., Pasolli, E., Fabbrini, M., Vitali, F., Agamennone, V., Dötsch, A., Selberherr, E., Grau, J. H., Meixner, M., Liere, K., Ercolini, D., de Filippo, C., Caderni, G., Brigidi, P., & Turroni, S. (2022). The Core Human Microbiome: Does It Exist and How Can We Find It? A Critical Review of the Concept. *Nutrients*, *14*(14), 2872. https://doi.org/10.3390/nu14142872

Stewart, C. J., Ajami, N. J., O'Brien, J. L., Hutchinson, D. S., Smith, D. P., Wong, M. C., Ross, M. C., Lloyd, R. E., Doddapaneni, H., Metcalf, G. A., Muzny, D., Gibbs, R. A., Vatanen, T., Huttenhower, C., Xavier, R. J., Rewers, M., Hagopian, W., Toppari, J., Ziegler, A.-G., … Petrosino, J. F. (2018). Temporal development of the gut microbiome in early childhood from the TEDDY study. *Nature*, *562*(7728), Article 7728. https://doi.org/10.1038/s41586-018-0617-x

Strachan, D. (2000). Family size, infection and atopy: The first decade of the “hygiene hypothesis.” *Thorax*, *55*(90001), 2S – 10. https://doi.org/10.1136/thorax.55.suppl_1.S2

Strachan, D. P. (1989). Hay fever, hygiene, and household size. *BMJ : British Medical Journal*, *299*(6710), 1259–1260.

Thompson, A. L., Monteagudo-Mera, A., Cadenas, M. B., Lampl, M. L., & Azcarate-Peril, M. A. (2015). Milk- and solid-feeding practices and daycare attendance are associated with differences in bacterial diversity, predominant communities, and metabolic and immune function of the infant gut microbiome. *Frontiers in Cellular and Infection Microbiology*, *5*. https://www.frontiersin.org/articles/10.3389/fcimb.2015.00003

Tun, H. M., Konya, T., Takaro, T. K., Brook, J. R., Chari, R., Field, C. J., Guttman, D. S., Becker, A. B., Mandhane, P. J., Turvey, S. E., Subbarao, P., Sears, M. R., Scott, J. A., Kozyrskyj, A. L., Sears, M. R., Subbarao, P., Anand, S. S., Azad, M., Becker, A. B., … the CHILD Study Investigators. (2017). Exposure to household furry pets influences the gut microbiota of infants at 3–4 months following various birth scenarios. *Microbiome*, *5*(1), 40. https://doi.org/10.1186/s40168-017-0254-x

Umesaki, Y., Setoyama, H., Matsumoto, S., & Okada, Y. (1993). Expansion of alpha beta T-cell receptor-bearing intestinal intraepithelial lymphocytes after microbial colonization in germ-free mice and its independence from thymus. *Immunology*, *79*(1), 32–37.

Walker, J. A., & McKenzie, A. N. J. (2018). TH2 cell development and function. *Nature Reviews Immunology*, *18*(2), Article 2. https://doi.org/10.1038/nri.2017.118

Wiley, K. S., Gregg, A. M., Fox, M. M., Lagishetty, V., Sandman, C. A., Jacobs, J. P., & Glynn, L. M. (2024). Contact with caregivers is associated with composition of the infant gastrointestinal microbiome in the first 6 months of life. *American Journal of Biological Anthropology*, *183*(4), e24858. https://doi.org/10.1002/ajpa.24858

Yasmin, F., Tun, H. M., Konya, T. B., Guttman, D. S., Chari, R. S., Field, C. J., Becker, A. B., Mandhane, P. J., Turvey, S. E., Subbarao, P., Sears, M. R., CHILD Study Investigators, Scott, J. A., Dinu, I., Kozyrskyj, A. L., Anand, S. S., Azad, M. B., Becker, A. B., Befus, A. D., … Turvey, S. E. (2017). Cesarean Section, Formula Feeding, and Infant Antibiotic Exposure: Separate and Combined Impacts on Gut Microbial Changes in Later Infancy. *Frontiers in Pediatrics*, *5*. https://www.frontiersin.org/articles/10.3389/fped.2017.00200

Yatsunenko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreras, M., Magris, M., Hidalgo, G., Baldassano, R. N., Anokhin, A. P., Heath, A. C., Warner, B., Reeder, J., Kuczynski, J., Caporaso, J. G., Lozupone, C. A., Lauber, C., Clemente, J. C., Knights, D., … Gordon, J. I. (2012). Human gut microbiome viewed across age and geography. *Nature*, *486*(7402), Article 7402. https://doi.org/10.1038/nature11053

Zheng, Y., Chaudhry, A., Kas, A., deRoos, P., Kim, J. M., Chu, T.-T., Corcoran, L., Treuting, P., Klein, U., & Rudensky, A. Y. (2009). Regulatory T-cell suppressor program co-opts

transcription factor IRF4 to control TH2 responses. *Nature*, *458*(7236), 351–356. https://doi.org/10.1038/nature07674